IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# Air Quality Forecasting And Early Warning System

1 Ast. Prof. Dr. Santosh Singh, 2 Sachin Manoj Chaurasiya, 3 Sumit Dinesh Singh, 1 Assistant Professor, Department of IT, Thakur College of Science and Commerce, Thakur Village, Kandivali (East), Mumbai, Maharashtra, India 2, 3 PG Student, Department of IT, Thakur College of Science and Commerce, Thakur Village, Kandivali (East), Mumbai, Maharashtra, India

**Abstract**

Delhi faces severe air pollution, especially during winter months when pollutant concentrations frequently exceed safe levels. The city's air quality is compromised by various factors, including vehicular emissions, industrial activities, dust from construction, and seasonal biomass burning. Real-time monitoring of pollutants like PM2.5, PM10, nitrogen dioxide ($NO_2$), sulfur dioxide ($SO_2$), and ammonia ($NH_3$) is critical to understanding the pollution patterns and developing forecasting models.

An air quality forecasting system uses historical and real-time data, combined with machine learning models, to predict pollution levels for future hours or days. This system incorporates variables such as traffic density, weather conditions, and emissions data to offer accurate forecasts of pollutant concentrations. The goal is to provide early warnings that allow authorities and citizens to take preventive measures.

This forecasting and early warning system is essential for mitigating health risks, particularly for vulnerable populations such as children, the elderly, and those with respiratory conditions. The data-driven predictions enable proactive actions, such as traffic restrictions, industrial activity limits, and advisories for staying indoors during high-pollution periods.

However, challenges persist in improving model accuracy and public compliance. Despite the availability of early warnings, public awareness and adherence to recommendations are often inadequate. Enhancing data integration from a wider range of sources and improving forecast resolution can increase the effectiveness of the system.

**Keywords:** Arima Model, Linear Regression Model, Air Quality Prediction, Air Quality Forecasting, Warning System, Air Quality Index(AQI), Data Extraction, PM2.5: Particulate matter smaller than 2.5 micrometers (in $\mu g/m^3$), PM10: Particulate matter smaller than 10 micrometers (in $\mu g/m^3$).

**Introduction**:

Air quality is a critical component of environmental health and public safety. Poor air quality can have immediate and severe impacts on human health, including respiratory issues, cardiovascular problems, and exacerbation of pre-existing conditions. In recent years, as urbanization and industrial activities have increased, so has the need for advanced systems to monitor and predict air quality to mitigate these effects.

An Air Quality Forecasting and Early Warning System (AQFEWS) is an integrated framework designed to provide timely and accurate information about air pollution levels. The goal of such a system is to forecast air quality trends, provide early warnings of potential air pollution events, and offer actionable insights to stakeholders, including public health officials, policymakers, and the general public.

Air quality forecasting and early warning systems (EWS) are critical in tackling air pollution in cities like Delhi, where air quality regularly deteriorates to hazardous levels, especially during the winter months. These systems provide timely information about upcoming air quality conditions, allowing authorities and the public to take preventive measures.

Air pollution is a major environmental issue with significant impacts on public health, the environment, and the economy. Rapid urbanization, industrial activities, and vehicular emissions contribute to deteriorating air quality, leading to increased respiratory diseases, cardiovascular conditions, and other health problems. Accurate forecasting and early warning systems are essential for mitigating these impacts by enabling timely interventions and public awareness.

**Literature Survey**

Air quality forecasting and early warning systems (AQFEWS) are designed to predict pollution levels and issue alerts to mitigate the impact of poor air quality on public health and the environment. These systems integrate data from various sources, employ predictive models, and communicate information to stakeholders.

Early systems primarily relied on observational data and simple statistical methods. Over time, advancements in satellite technology, sensor networks, and computational models have significantly enhanced the capabilities of AQFEWS.

- Historical Developments: Initially, air quality monitoring focused on real-time measurements with limited forecasting. Early models used linear regression and basic statistical approaches (e.g., the work of Seinfeld and Pandis, 2006).
- Technological Evolution: The advent of remote sensing technologies and the expansion of ground-based sensor networks have greatly improved data collection and forecasting capabilities (e.g., Nowak et al., 2006).

Effective AQFEWS rely on diverse data sources, including:

- Ground-based Sensors: Provide high-resolution, localized data on various pollutants (e.g., PM2.5, NO2, O3).
- Satellite Observations: Offer broad spatial coverage and valuable data on aerosol optical depth and other atmospheric parameters (e.g., Martin et al., 2002).
- Meteorological Data: Essential for understanding the dispersion and transformation of pollutants (e.g., Baklanov et al., 2014).

Environmental pollution has been one of the most serious scourges because of the great damage it causes to economies and the lives of people. According to the World Health Organization (WHO) report from March 6, 2017, air pollution takes the lives of 1.7 million children under 5 years of age every year (http://www.who.int/mediacentre/news/releases/2017/pollution-child-death/en/).

With industrialization and urbanization, air pollution and hazy weather have increased at a great rate, especially in developing countries (Wu and Zhang, 2017). In recent years, China has experienced severe and persistent air pollution in the winter, which has attracted worldwide attention (Liu et al., 2017). Relevant research about this issue has also flourished (Ma et al., 2017; Zhuo et al., 2017).

The dominant pollution contaminants, such as particulate matter (PM), sulfur dioxide ($SO_2$) and nitrogen oxides ($NO_x$), pose significant risks to environments. These pollution contaminants are barometers of reaction to environmental problems (L. Zhang et al., 2017). In most air quality monitoring systems, carbon monoxide (CO) and ozone ($O_3$) are the two main objective pollution contaminants (Maga et al., 2017). Additionally, other hazardous substances, such as radiation, soiling or specific toxic gases, cause great damage to air quality. In fact, finding the main factor that causes air pollution is the key to solving the problem. Moreover, precise prediction of pollution contaminants plays a vital role. Therefore, it is of great importance to propose an early-warning system and take effective corresponding protection measures.

There are a series of environmental factors that cause air pollution, and the flow and diffusion of pollution contaminants is a rather complex process (Vidale et al., 2017). In early stages, many scholars focused on the relationship between air pollution and its environmental influence factors (Pahlavani et al., 2017). Based on these pioneering works, air quality monitoring systems emerged for monitoring the main pollution contaminants. Recently, studies on the emission sources of pollution contaminants have become a hot topic, and early-warning systems have shown practical importance (Yin et al., 2015; Liora et al., 2016; L. Li et al., 2017). However, little research has focused on the selection of pollutant indicators and the design of early-warning systems for particular cities.

On the other hand, a variety of models have been proposed to forecast pollution contaminants that can be roughly divided into two categories: statistical models and machine learning models. Traditional statistical approaches, including regression models (Lee et al., 2017), time series models (Nhung et al., 2017) and autoregressive moving average models (ARIMA) (Zafra et al., 2017), are applied in pollution contaminant prediction. Additionally, researchers have investigated the partial and multiple linear features of pollution contaminants and have shown different patterns of statistical models (Carlsen et al., 2018; Zhu et al., 2017). However, each of these approaches has difficulties in dealing with non-linear problems.

**Methodology:**

**1. Data Collection and Integration**

• Monitoring Network: Expand and optimize the existing air quality monitoring stations across Delhi to ensure comprehensive coverage. The stations should measure key pollutants like $PM_{2.5}$, $PM_{10}$, $NO_2$, $SO_2$, CO, $O_3$, etc.

• Satellite Data: Use satellite data to fill gaps where ground stations are sparse or non-existent. Satellite-based remote sensing can monitor larger regions and capture trends beyond the city.

• Mobile Sensors: Utilize mobile air quality sensors in public transport (buses, taxis) and drone-based monitoring for hotspot identification.

• Crowdsourced Data: Leverage crowdsourcing tools where citizens can report visible pollution issues or personal monitoring device data.

• Weather Data Integration: Integrate weather data (e.g., temperature, humidity, wind speed, and direction) from meteorological stations to improve accuracy in forecasting.

**2. Air Quality Forecasting Models**

• Predictive Modeling: Develop or adapt predictive models (e.g., machine learning or statistical models) that can forecast air quality based on historical data and real-time inputs. Include models that can predict up to 72 hours in advance.

• Source Identification: Incorporate emissions inventories and models to identify and quantify pollution sources such as traffic, industry, biomass burning, and seasonal factors like Diwali fireworks and crop stubble burning.

• Machine Learning/AI: Employ machine learning techniques to improve the accuracy of forecasts by identifying patterns in historical and real-time data.

• Scenario-based Forecasting: Introduce what-if scenarios in forecasting, such as the impact of traffic control, industrial shutdowns, or crop burning on air quality, to inform policy decisions.

**3. Early Warning and Public Dissemination**

• User-friendly Platforms: Create a centralized digital platform (web/mobile app) that disseminates real-time air quality data and forecasts. The platform should be accessible and easily understandable by the general public.

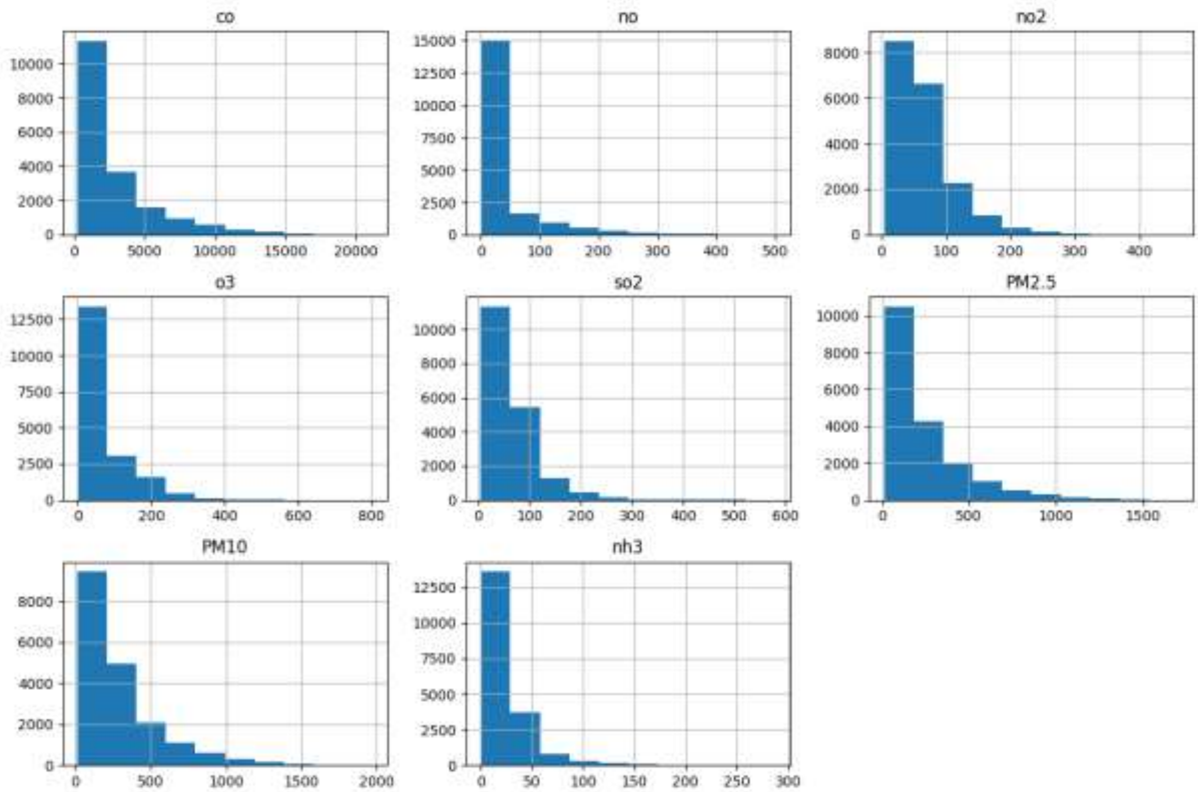
• Air Quality Index (AQI): Use a clear AQI system with color codes and health advisories based on forecasted air quality levels. Provide guidance on protective actions (e.g., avoiding outdoor activities, wearing masks).

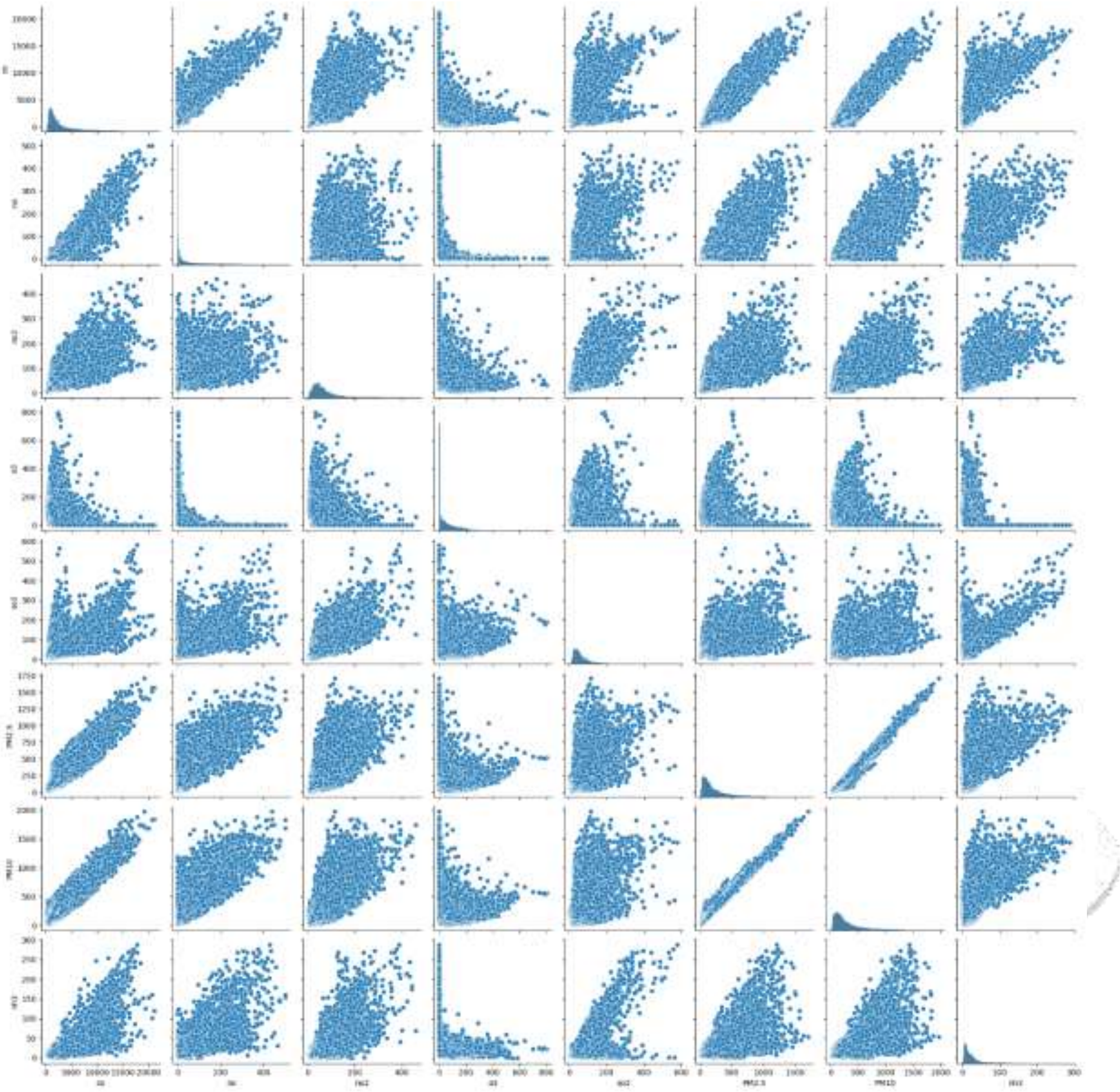
• Push Notifications: Enable early warning systems to send notifications or alerts to the public based on forecasted spikes in pollution. This can be done via SMS, app notifications, or other popular communication channels.

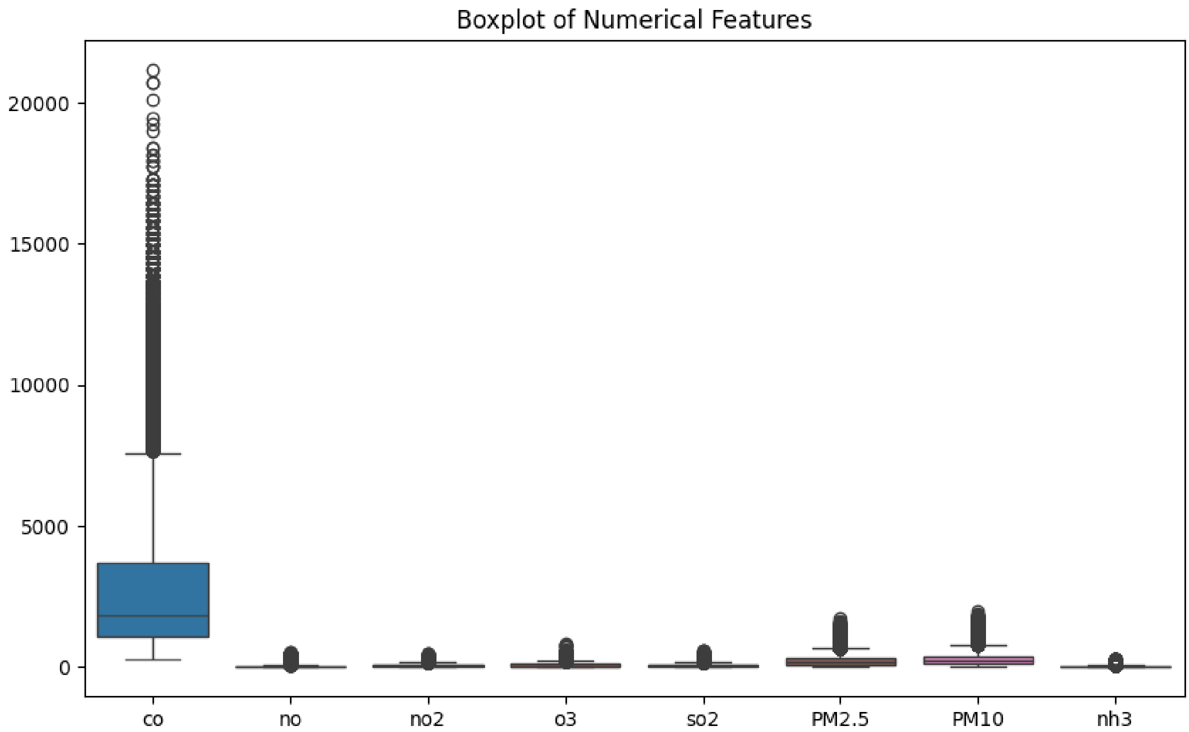
• Integration with Smart Cities: Collaborate with existing smart city initiatives to integrate air quality data with urban planning, traffic control systems, and public health responses.

• Social Media Integration: Regular updates on air quality forecasts and warnings should be shared via social media channels, targeting various demographics.

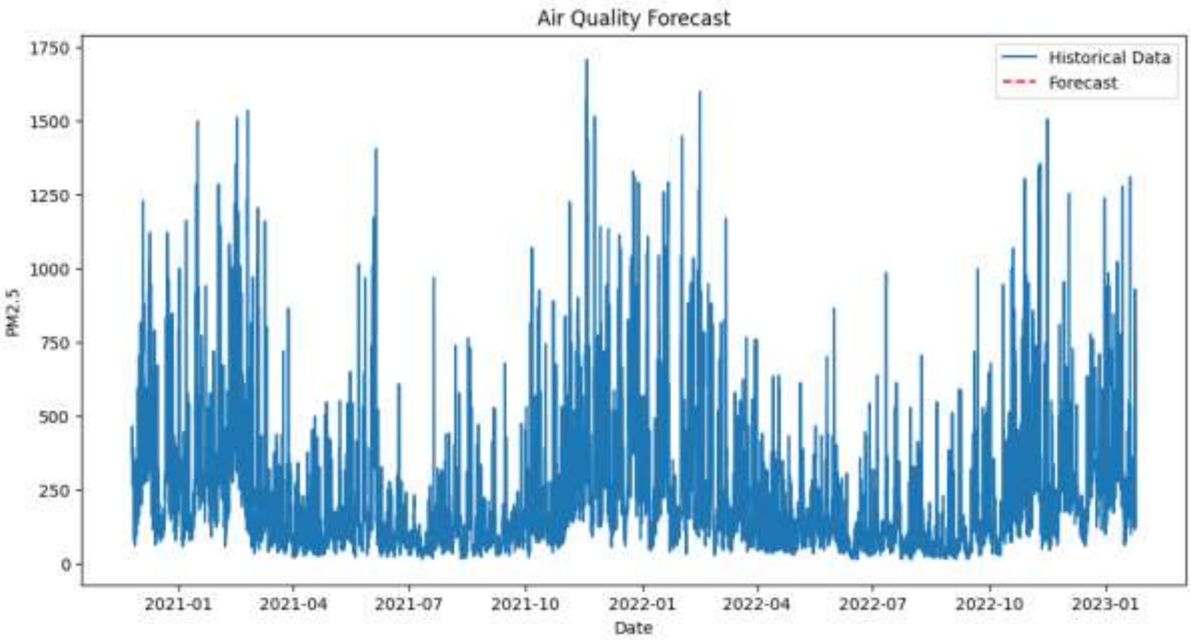
EDA





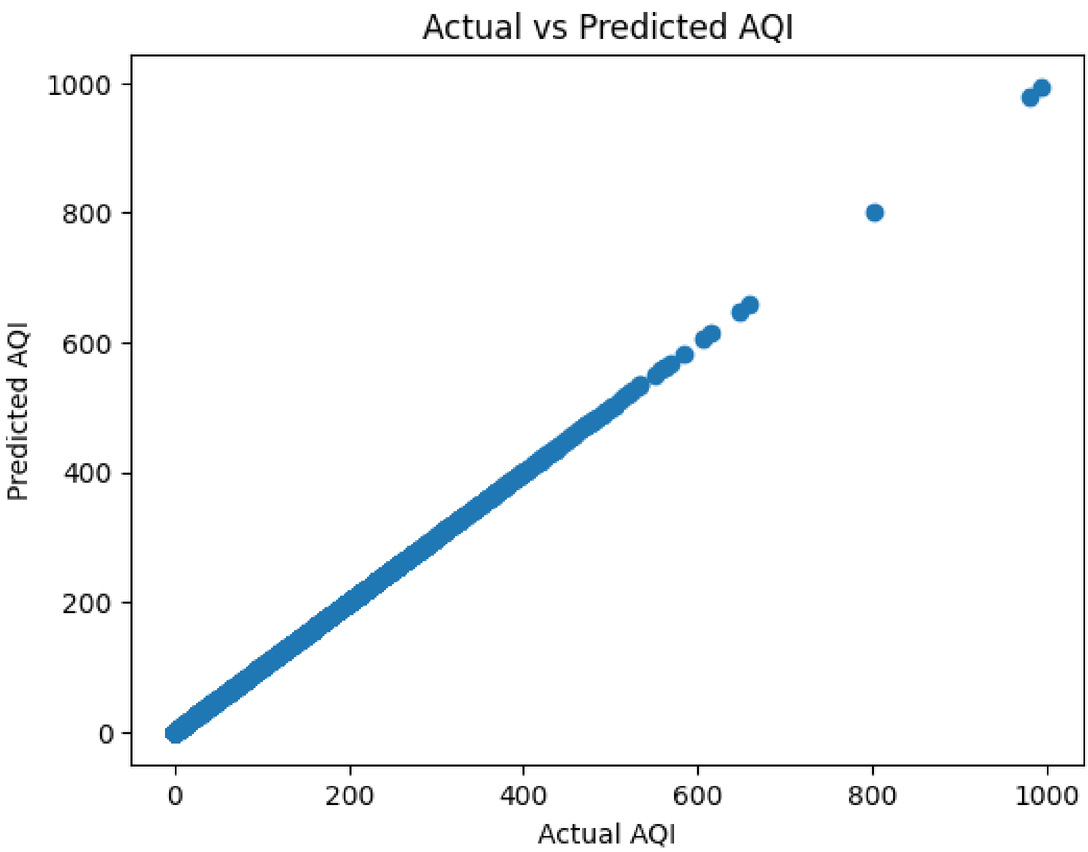

Boxplot of Numerical Features
20000
15000
10000
5000
0
co
no
no2
o3
so2
PM2.5
PM10
nh3

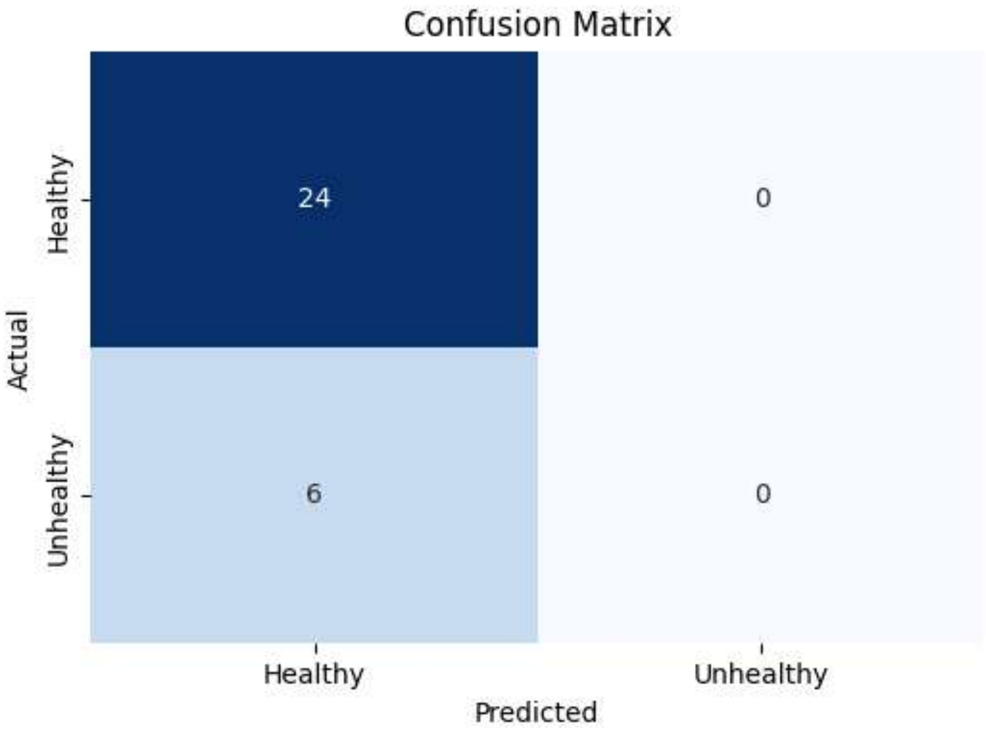
Results




Accuracy: 80.00%

**Actual vs Predicted AQI(Air Quality Index)**






Mean Absolute Error (MAE): 253.057440415009

Root Mean Squared Error (RMSE): 359.9098807653383

Mean Absolute Percentage Error (MAPE): nan%

Air quality forecast is within safe limits.

Expected Outcomes:

1. Accurate Forecasting: A reliable forecasting model that predicts air quality with high precision, allowing stakeholders to anticipate pollution events and plan accordingly.
2. Effective Early Warning: A responsive early warning system that provides timely alerts and actionable information to mitigate the effects of poor air quality.
3. Informed Public and Policy Actions: Enhanced public awareness and informed decision-making by authorities, leading to better health outcomes and more effective air quality management.

**Conclusion:**

Therefore, this paper presented a highly efficient machine learning model combining Convolutional Neural Networks (CNN), Random Forest (RF), and Support Vector Machine (SVM) to differentiate between healthy and unhealthy nails. Such an accuracy in the detection of nail diseases is achieved by this model because feature extraction gets done by CNN; and the initial classification gets done by RF, while the results get refined by SVM. This solution shall help scale early disease detection, especially for countries like India that have less access to health care, thus enabling the proper identification and treatment of nail conditions at the proper time.

Although this model seems to be promising for encouraging outcomes, further refinements can be made by extending the size of the dataset and adding some sophisticated techniques. Having it implemented into real-life scenarios and integrated with applications on mobile devices will make it more accessible-especially in areas of remote places. This work leads to setting a premise for health solutions reliant on AI technologies that enhance early detection of diseases in resource-constrained settings.

## REFERENCES

Baklanov, A., et al. (2014). "Modeling of atmospheric composition: A review of current practices and research directions." Atmospheric Chemistry and Physics.

DonnellyA. et al. “Real time air quality forecasting using integrated parametric and non-parametric regression techniques”

CobournW.G. “An enhanced PM2.5 air quality forecast model based on nonlinear regression and back-trajectory concentrations”

BaiY. et al. “Air pollutants concentrations forecasting using back propagation neural network based on wavelet decomposition with meteorological conditions”

Cohan, D. S., et al. (2006). "Improved forecast of PM2.5 using the Community Multiscale Air Quality (CMAQ) model." Journal of Air & Waste Management Association.

Dowell, M. D., et al. (2011). "Real-time air quality forecasting: An overview of current methods and emerging technologies." Environmental Monitoring and Assessment.